ETH zürich

Optimization in Energy Systems

Newton Raphson, Interior Point

Prof. Gabriela Hug, ghug@ethz.ch

Learning Objectives

- After this lecture, you should be able
  - to describe in three sentences how the Newton Raphson method works
  - to apply the Newton Raphson method to a nonlinear equation system
  - to describe in three sentences how the Interior Point method works
  - to apply the Interior Point method (combined with Newton-Raphson) to a non-linear optimization problem

Optimality Conditions

- Lagrange Function

$$L = f(x) + \lambda^T \cdot h(x) + \mu^T \cdot g(x)$$

- Minimize Lagrange Function

$$\frac{\partial L}{\partial x} = \frac{\partial}{\partial x}\left(f(x) + \lambda^T h(x) + \mu^T g(x)\right)\Big|_{\hat{x},\hat{\lambda},\hat{\mu}} = 0$$

$$\frac{\partial L}{\partial \lambda} = h(x)\,|_{\hat{x}} = 0$$

$$\frac{\partial L}{\partial \mu} = g(x)\,|_{\hat{x}} \leq 0$$

$$diag\{\mu\}\frac{\partial L}{\partial \mu} = diag\{\mu\}\, g(x)\,|_{\hat{x},\hat{\mu}} = 0$$

$$\mu \geq 0$$

$\Rightarrow$ Karush-Kuhn Tucker first order or necessary optimality conditions

Newton-Raphson Method

- One-dimensdional: Find solution to

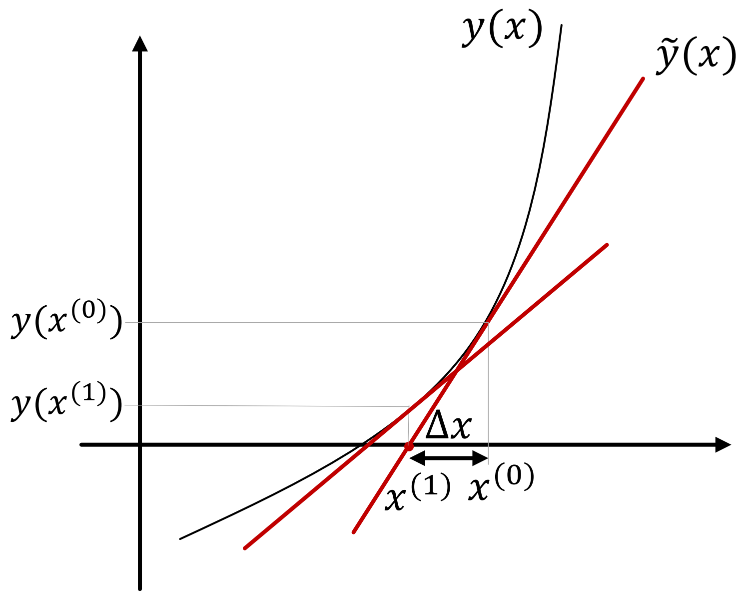
$$y(x) = 0$$

- Taylor's Theorem

$$\tilde{y}(x + \Delta x) = y(x) + \frac{dy}{dx}\Delta x$$

$$\tilde{y}\left(x^{(0)} + \Delta x\right) = y\left(x^{(0)}\right) + \left.\frac{dy}{dx}\right|_{x=x^{(0)}} \cdot \Delta x = 0$$

$$\Rightarrow \Delta x = -\left(\left.\frac{dy}{dx}\right|_{x=x^{(0)}}\right)^{-1} \cdot y(x^{(0)})$$



Newton-Raphson Method

- Two-dimensional
  - Equations

$$y_1(x_1, x_2) = 0$$
$$y_2(x_1, x_2) = 0$$

  - Taylor Approximation

$$y_1(x_1 + \Delta x_1, x_2 + \Delta x_2) = y_1(x_1, x_2) + \frac{\partial y_1}{\partial x_1}\Delta x_1 + \frac{\partial y_1}{\partial x_2}\Delta x_2 = 0$$

$$y_2(x_1 + \Delta x_1, x_2 + \Delta x_2) = y_2(x_1, x_2) + \frac{\partial y_2}{\partial x_1}\Delta x_1 + \frac{\partial y_2}{\partial x_2}\Delta x_2 = 0$$

$$\Rightarrow \quad \underbrace{\begin{bmatrix} y_1(x_1, x_2) \\ y_2(x_1, x_2) \end{bmatrix}}_{y(x)} + \underbrace{\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix}}_{J(x^{(\nu)})} \cdot \underbrace{\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}}_{\Delta x} = 0$$

$$\Rightarrow \quad \Delta x = -J\left(x^{(\nu)}\right)^{-1} \cdot y(x^{(\nu)})$$

Newton-Raphson Method

- Multi-dimensional Case

$$\boldsymbol{y}(\boldsymbol{x}) = [y_1(\boldsymbol{x}), y_2(\boldsymbol{x}), \ldots, y_n(\boldsymbol{x})]^T = 0$$

$$\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$$

- Taylor Approximation

$$\boldsymbol{y}\left(\boldsymbol{x}^{(\nu)} + \Delta\boldsymbol{x}^{(\nu)}\right) \approx \boldsymbol{y}\left(\boldsymbol{x}^{(\nu)}\right) + \boldsymbol{J}\left(\boldsymbol{x}^{(\nu)}\right)\Delta\boldsymbol{x}^{(\nu)} = 0$$

$$J = \frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}$$

$\Rightarrow$ solve for $\Delta\mathbf{x}^{(\nu)}$

Newton-Raphson Method

1. Set $\nu = 0$ and choose an appropriate starting value $x^{(0)}$.
2. Compute $y\left(x^{(\nu)}\right)$.
3. Compare $y\left(x^{(\nu)}\right)$ with specified tolerance ϵ
   a. If $\left|y\left(x^{(\nu)}\right)\right| < \epsilon$: solution found, stop.
   b. Otherwise: go to next step.
4. Compute $J(x^{(\nu)})$.
5. Compute $\Delta x = -J\left(x^{(\nu)}\right)^{-1} \cdot y(x^{(\nu)})$.
6. Update $x^{(\nu+1)} = x^{(\nu)} + \Delta x$
7. Update iteration counter, i.e. $\nu \rightarrow \nu + 1$.

Solving Optimization Problem

- Interior Point Method
  - Add barrier term to objective function

$$\min_x f(x) \quad s.t.\ g(x) \leq 0 \quad \Longrightarrow \quad \min_x f(x) + \zeta \cdot B(g(x))$$

barrier function

$\Rightarrow$ Barrier term avoids leaving feasible region

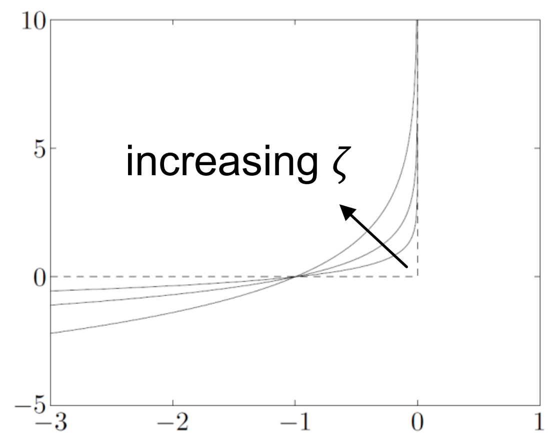
  - Ideal properties of barrier term

$$B(x) = \begin{cases} 0, & g(x) \leq 0 \\ \infty, & g(x) > 0 \end{cases}$$

$\Rightarrow$ approximation

$$\zeta \cdot B\big(g(x)\big) = -\zeta \ln(-g(x))$$

=> twice differentiable



Source: “Convex Optimization” by Stephen Boyd, Lieven Vandenberghe

Solving Optimization Problem

- Interior Point Method
  - Discussion of barrier term
    - ζ large: approximation of barrier term not the greatest
    - ζ small: approximation improves but the function $\zeta \cdot B(g(x))$ difficult to minimize by Newton's method because hessian varies rapidly near boundary of feasible set

=> use iterative procedure letting ζ go towards smaller values; should ideally become zero at the end because otherwise suboptimal solution

Solving Optimization Problem

- Interior Point Method
  - Reformulation of the optimization problem

$$\begin{aligned} &\min_x f(x) \\ &s.t.\ h(x) = 0 \\ &\qquad g(x) \leq 0 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} &\min_x f(x) - \zeta \sum_{i=1}^{p} \ln(s_i) \qquad (\zeta > 0) \\ &s.t.\ h(x) = 0 \\ &\qquad g(x) + s = 0 \\ &\qquad s > 0 \end{aligned}$$

barrier function

$\Rightarrow$ barrier parameter ζ has to become (almost) zero in optimum

$\Rightarrow$ barrier function implicitly assumes $s > 0$, i.e. it acts as a barrier to not cross the border at zero

- Lagrange Function

$$L_{IP} = f(x) - \zeta \sum_{i=1}^{p} \ln(s_i) + \lambda^T h(x) + \mu^T (g(x) + s)$$

Solving Optimization Problem

$$L_{IP} = f(x) - \zeta \sum_{i=1}^{p} \ln(s_i) + \lambda^T h(x) + \mu^T (g(x) + s)$$

- Interior Point Optimality Conditions

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x} &= \frac{\partial}{\partial x}\left(f(x) + \lambda^T h(x) + \mu^T g(x)\right)|_{\hat{x},\hat{\lambda},\hat{\mu}} &= 0 \qquad & y_x(x,\lambda,\mu,s) = 0\\
\frac{\partial \mathcal{L}}{\partial \lambda} &= h\,(x)|_{\hat{x}} &= 0 \qquad & y_\lambda(x,\lambda,\mu,s) = 0\\
\frac{\partial \mathcal{L}}{\partial \mu} &= g\,(x)|_{\hat{x}} + \hat{s} &= 0 \qquad & y_\mu(x,\lambda,\mu,s) = 0\\
\frac{\partial \mathcal{L}}{\partial s} &= \mu - \zeta_k \cdot diag\{\frac{1}{s_i}\} \cdot I|_{\hat{x},\hat{\mu},\hat{s}} &= 0 \qquad & y_s(x,\lambda,\mu,s) = 0\\
& \hat{\mu} &\geq 0 & \\
& \hat{s} &> 0 &
\end{aligned}$$

- Application of Newton-Raphson Method

$$\underbrace{\begin{pmatrix} H_{\mathcal{L}} & \nabla h^{\mathrm{T}} & \nabla g^{\mathrm{T}} & 0 \\ \nabla h & 0 & 0 & 0 \\ \nabla g & 0 & 0 & I \\ 0 & 0 & \mathrm{diag}(s) & \mathrm{diag}(\mu) \end{pmatrix}}_{J(x^{(k)},\lambda^{(k)},\mu^{(k)},s^{(k)})} \cdot \underbrace{\begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta \mu \\ \Delta s \end{pmatrix}}_{\Delta^{(k)}} = -\underbrace{\begin{pmatrix} \nabla \mathcal{L} \\ h \\ g + s \\ \mathrm{diag}(\mu) \cdot s - \zeta_k \end{pmatrix}}_{d(x^{(k)},\lambda^{(k)},\mu^{(k)},s^{(k)})}.$$

Solving Optimization Problem

- Solution Algorithm

1. Choose initial values $x^{(0)}, \lambda^{(0)}, \mu^{(0)}, s^{(0)}$ and ζ_k and set iteration counter $k = 0$.
2. Build the Jacobian matrix $J^{(k)}$ and the vector $d^{(k)}$.
3. Determine the values for the update vector $\Delta^{(k)}$ by solving the linear equation system $J^{(k)} \cdot \Delta^{(k)} = d^{(k)}$.
4. Update all variables, i.e.

$$
\begin{aligned}
x^{(k+1)} &= x^{(k)} + \kappa_s \cdot \Delta x^{(k)} \\
\lambda^{(k+1)} &= \lambda^{(k)} + \kappa_\mu \cdot \Delta \lambda^{(k)} \\
\mu^{(k+1)} &= \mu^{(k)} + \kappa_\mu \cdot \Delta \mu^{(k)} \\
s^{(k+1)} &= s^{(k)} + \kappa_s \cdot \Delta s^{(k)}
\end{aligned}
$$

$$
\begin{aligned}
\kappa'_s &= \min_{\Delta s_i^{(k)} < 0} \left\{ \frac{-s_i^{(k)}}{\Delta s_i^{(k)}} \right\} \\
\kappa_s &= \min\{1, 0.9995 \cdot \kappa'_s\} \\
\kappa'_\mu &= \min_{\Delta \mu_i^{(k)} < 0} \left\{ \frac{-\mu_i^{(k)}}{\Delta \mu_i^{(k)}} \right\} \\
\kappa_\mu &= \min\{1, 0.9995 \cdot \kappa'_\mu\}
\end{aligned}
$$

5. Update the barrier parameter ζ_k as follows

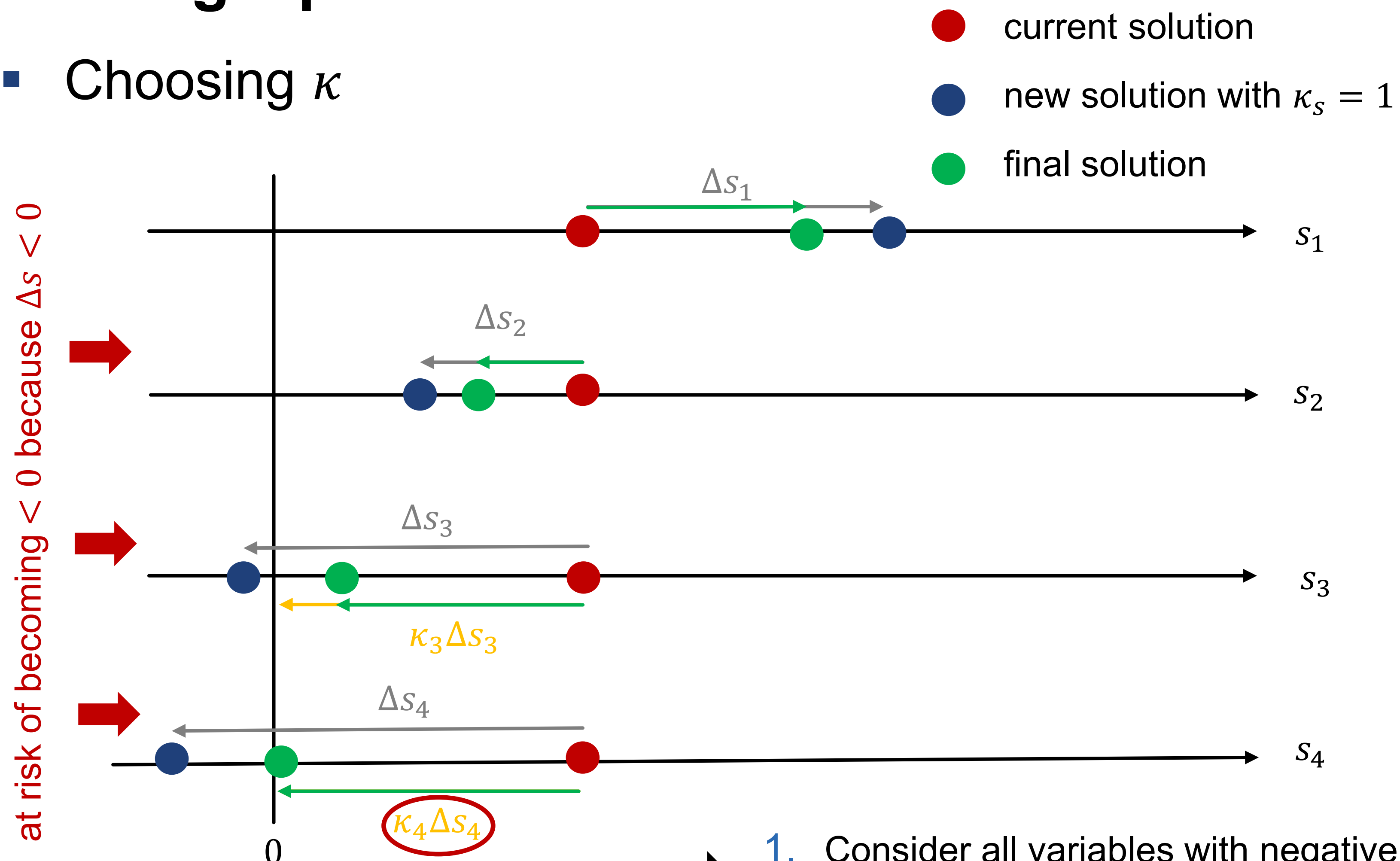
$$\zeta^{(k+1)} = \beta \cdot \frac{\mu^{(k)^T} s^{(k)}}{2m}$$

where β is a tuning parameter which is usually in the range of 0.1 … 0.5 and m is the number of unknowns.

6. If all values in $\Delta^{(k)}$ are smaller than a pre-defined tolerance, convergence is reached. Otherwise, increase the iteration counter and go to step 2.

Solving Optimization Problem

- Choosing κ



1. Consider all variables with negative Δ
2. Identify variable that requires smallest κ

Solving Optimization Problem

- Example

$$\min_{x_1, x_2} (x_1^2 + x_2^2)$$

$$s.t. \; -x_1 - x_2 + 4 \leq 0$$

Summary

- Newton Raphson Method
  - to solve $y(x) = 0$, iteratively use $\Delta x = -J^{-1}\left(x^{(\nu)}\right) \cdot y\left(x^{(\nu)}\right)$

- Interior Point
  - to solve inequality constrained optimization problem

$$\min_{x} f(x) - \zeta \sum_{i=1}^{p} \ln(s_i) \qquad (\zeta > 0)$$

barrier function

$$s.t.\ h(x) = 0$$

$$g(x) + s = 0$$

$$s > 0$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x} &= \frac{\partial}{\partial x}\left(f(x) + \lambda^T h(x) + \mu^T g(x)\right)|_{\hat{x},\hat{\lambda},\hat{\mu}} &= 0\\
\frac{\partial \mathcal{L}}{\partial \lambda} &= h(x)|_{\hat{x}} &= 0\\
\frac{\partial \mathcal{L}}{\partial \mu} &= g(x)|_{\hat{x}} + \hat{s} &= 0\\
\frac{\partial \mathcal{L}}{\partial s} &= \mu - \zeta_k \cdot diag\{\frac{1}{s_i}\} \cdot I|_{\hat{x},\hat{\mu},\hat{s}} &= 0\\
& \hat{\mu} &\geq 0\\
& \hat{s} &\geq 0
\end{aligned}$$